SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

AD-A231 380

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

JAN 22 1991

1a REPORT SECURITY CLASSIFICATION
Unclassified

1b RESTRICTIVE MARKINGS

2a SECURITY CLASSIFICATION AUTHORITY

2b DECLASSIFICATION/DOWNGRADING SCHEDULE

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
N00014-85-K-0423

5. MONITORING ORGANIZATION REPORT NUMBER(S)

6a. NAME OF PERFORMING ORGANIZATION
Carnegie Mellon Univ.

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
Personnel & Training

6c. ADDRESS (City, State, and ZIP Code)
Pittsburg, PA. 15213

7b. ADDRESS (City, State, and ZIP Code)
Code 442PT
700 N. Quincy St. Arlington, VA. 22217

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
Office of the Chief of Naval Research

8b. OFFICE SYMBOL (If applicable)

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
RT #442c005---02-12-87(1142cs)

8c. ADDRESS (City, State, and ZIP Code)
800 N. Quincy St.
Arlington, VA. 22217-5000
ATTN: Code 1511B:DRVM

10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Expert Planning Processes in Writing (Unclassified)

12. PERSONAL AUTHOR(S)
John R. Hayes & Linda S. Flower

13a. TYPE OF REPORT Final
13b. TIME COVERED FROM 6/15/85 TO 6/14/8
14. DATE OF REPORT (Year, Month, Day) 90/12/17
15. PAGE COUNT 52

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report summarizes research conducted under contract No. 00014-85-K-0423 on the nature of planning processes in writing. Section 1 presents a general characterization of planning based on an integration of planning research in the fields of A. I., cognitive science, and writing. This characterization provides a framework for studying planning in writing. Section 2 summarizes two protocol studies designed to identify characteristics of planning in writing. Several differences among expert and novice writing strategies are identified. Section 3 reports two experimental studies of skills which are fundamental to planning in writing. The first explores writer's ability to judge when the use of metaphor will help readers to comprehend a text. The second demonstrates that providing writers with topic knowledge can make them insensitive to the readers need to have that topic knowledge explained to them.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION

22a. NAME OF RESPONSIBLE INDIVIDUAL
John R. Hayes

22b. TELEPHONE (Include Area Code)
412-268-2803

22c. OFFICE SYMBOL

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

90 12 27 075

ONR FINAL REPORT

EXPERT PLANNING PROCESSES IN WRITING

JOHN R. HAYES and LINDA S. FLOWER

Projects of ONR work appearing in the Technical Reports of the Center for the Study of Writing, Carnegie Mellon University, Pittsburgh, PA

32. Foundations for Creativity in the Writing Process: Rhetorical Representations of Ill-Defined Problems — L.J. Carey & L. Flower

    In J.A. Glover, R.R. Ronning, & C.R. Reynolds (Eds.), <u>Handbook of Creativity</u>. New York: Plenum Press. (1989)

34. Planning in Writing: The Cognition of a Constructive Process+ — L. Flower, K.A. Schriver, L. Carey, C. Haas, & J.R. Hayes

35. Differences in Writers' Initial Task Representations+ — L. Carey, L. Flower, J.R. Hayes, K.A. Schriver, & C. Haas

Final Report: Expert Planning Processes in Writing

John R. Hayes and Linda S. Flower

This is a final report of research accomplished under a grant to study "Expert planning processes in writing" from the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval research, under Contract No. 00014-85-K-0423. The report is divided into three sections. The first section reports a review and synthesis of planning research in the fields of artificial intelligence, industrial design, and writing titled "On the Nature of Planning in Writing" by John R. Hayes. The second section reports two observational studies on planning processes in writing: "Differences in writers'initial task representation" by Linda Carey, Linda Flower, John R. Hayes, Karen A. Schriver, and Christina Haas and "Planning in writing: The cognition of a constructive process". The third section, reports two experimental studies on skills which are fundamental to planning in writing: "Designing Instructional Texts: Metaphor as Explanatory Strategy." by Lorraine Higgins, John R. Hayes, & Rebecca Burnett and "The Effect of Topic Knowledge in Editing." by John R. Hayes, Karen A Schriver, Andrea Blaustein, & Rachel Spilka. The studies in section two are already published and are, therefore, described only briefly. The studies in sections 1 and 3 have not as yet been published and are described in greater detail.

Section 1: On the Nature of Planning in Writing

Planning is widely recognized among writing researchers as a very important skill for successful writing. Carey, Flower, Hayes, Schriver, and Haas (1987) have found that more successful writers plan more completely than do less successful writers. Bereiter, Scardamalia, and their colleagues (Bereiter & Scardamalia, 1987: Scardamalia, Bereiter, & Steinbach ,1984) have conducted a number of interesting experiments to help novice writers improve their planning skills. Despite this interest, the theory of planning is much less well developed in the field of writing research than it is in the areas of cognitive science and artificial intelligence. My purpose in this essay is to review some of the ideas about planning developed in cognitive science and artificial intelligence and to attempt to integrate those ideas

into a coherent view which reveals their implications for writing research. The essay consists of seven sections on the following topics:

1.) The nature of planning,
2.) Why planning is useful,
3.) Types of planning,
4.) Multiple planning spaces, layering, and meta-planning,
5.) Features of the task and task environment which influence planning, and
6.) Summary

The Nature of Planning

Suppose that we are faced with a task. We have an outcome in mind but don't know as yet what action or sequence of actions to take to get achieve that outcome. If the actions available to us are expensive either in time or materials then we are likely to engage in planning before taking action. That is, we are likely to try to foresee the effects of the actions before we actually perform them. In common sense terms, planning is a process of looking before we leap. We may do this by imagining actions, by drawing sketches, or by doing other things that symbolize the actions. Chess players review alternatives silently before committing themselves to a move on the board. Carpenters sketch the framing of a house on paper before they begin to cut and nail lumber. Sculptors will make a model in malleable clay before casting a work in permanent bronze. In all of these cases, the planning environment is very different from the action environment: Unobservable thought vs. an observable move on the chess board; erasable marks on paper vs. wood irreversibly cut; inexpensive clay vs. costly bronze.

I have tried to capture the relation between planning and action in Figure 1. This account is a slightly modified version of Newell and Simon's (1972, p429) description of planning. I will discuss the differences below in the section on types of planning. The **action space** is the world of the original task. In a carpentry task, the action space encompasses the object to be built and the materials and tools to be used in building it. A **planning space** is a separate space in which an image of the task can be created. A planning space for the carpentry might be a world of pencil and paper in which the object to be built and the forms from which it is to be built can be drawn. An alternative planning space might be a world of mental imagery.

Steps in planning

The process of planning may be thought of as involving three steps. First, an image of the task is created in the planning space. This image

typically includes a representation of the desired outcome, or goal, of the task. The image of the goal may be quite incomplete. For example, in an industrial design task, it may consist of a list of criteria for the finished product such as size, appearance, cost, etc. The image of the task may also include a representation of the current situation, e.g., drawings of parts to be used, of the methods available for producing the desired outcome, e.g., manufacturing procedures such as injection molding, and of constraints or limitation under which the task is to be carried out, e.g., delivery dates.

The second step in the planning process is to carry out the image of the task in the planning space. In an industrial design task, this might involve making sketches of alternative configurations for the parts, choosing one configuration, and making detailed drawings of it. It is important to note that the account of planning presented here allows the second planning step to produce two rather different products. One of them consists of further specification of the goal, e.g., choices of color, materials, etc. The second is a **plan**, that is, a specified sequence of actions to be taken or subgoals to be accomplished in carrying out the task. This characterization of planning is consistent with accounts of planning in writing in Hayes and Flower (1980) and in Flower and Hayes (1984).

In some tasks, goal specification seems to be more important. For example, in architectural and industrial design tasks, the primary objective of planning is to specify the final form of the product rather than laying out the steps by which the product will be produced. In other tasks, the plan seems most important. For example, in building and manufacturing tasks, the primary objective of planning is to specify a sequence of steps to produce the product. In still other tasks such as writing, both seem to be important. The writer often needs to specify both what effects he or she hopes to have on the audience as well as a set of actions to produce that effect.

The final step in the planning process is to use the results of planning in carrying out the original task in the action space. This involves translating the new goal specifications into the action space and using the plan to guide the accomplishment of the task. Miller, Galanter, and Pribram (1960) describe plans as control structures, that is, as structures which convey information to control some other process. Control may be "tight" or "loose". An example of a very tight control structure is a computer program. The control structure is very tight in the sense that the computer performs just those actions specified by the program and no others. In interpreting the "control structure" metaphor, though, it is important to understand that most plans are loose rather than tight control structures. In most cases, it seems more natural to say that plans guide rather than control action.

Plans often underspecify actions. For example, in our study of the relation of planning to the writing of sentences (Kaufer, Hayes, and Flower 1986), we found that the plan might specify all of the topics to be discussed but that even the most complete plan did not specify more than a few of the words that the writer should use to express the topics. Choosing those words was in fact one of the major tasks on the way to turning the plan into action.

Typically, then, the plan is not fully in control of action. Rather, plans provide suggestions for action -- suggestions which may be accepted, rejected or modified as action proceeds. Certainly, writing plans are frequently modified in the course of execution. Topics may be added, deleted, combined, or otherwise modified as the writer tries to translate the initial plan into prose. Indeed, the activity of writing sometimes suggests new ideas which lead to radical changes in the writing plan.

While plans are control structures, not all control structures are plans. If we are to accept the common sense notion that plans are the result of planning and that planning is an intentional problem solving activity, then, many control structures are not plans. For example, the genetic structures which control instinctive behaviors such as migration in birds are not plans because they did not result from an intentional problem solving (planning) process.

Why Planning Is Useful?

Planning can be useful in two rather different ways. First, it can be useful by providing economy in the execution of the task. The cost of exploring is typically less in the planning space than in the action space. Thus, a sculptor may explore an esthetic problem in the planning space by drawing sketch after sketch, testing and rejecting one possibility after another before beginning to sculpt. Carrying out the same exploration in the action space, that is, by carving stone, would clearly be very expensive of time and materials.

The second way planning can be useful is by providing flexibility in the choice of problem solving strategies. Some problem solving strategies may be available in the planning space that are not available in the action space. For example, in planning a chess move, we can imagine a number of alternative moves that we might make and our opponent's potential replies to them to see which might give us the greatest advantage. In actual play, however, trying alternative moves on the board is illegal. But even if it

weren't illegal, it would be ineffective because it would reveal our game plan to the opponent. Thus, there may be constraints in the action space which prevent us from applying some problem solving methods. Methods which are often available in the planning space but not in the action space include working backward, hypothetical reasoning, and abstraction.

Kinds of Planning

Different kinds of planning can be distinguished on the basis of the most salient activity which the planning involves. Below, I discuss three frequently employed kinds of planning: planning by abstraction, planning by analogy, and planning by modeling.

Planning by abstraction

One of the most commonly discussed kinds of planning, planning by abstraction, might be characterized as planning in which only the most important or critical aspects of the problem are represented in the planning space. The difference between Newell and Simon's (1972) characterization of planning and the one presented here concerns abstraction. For Newell and Simon, the first step in planning is ". . . abstracting by omitting certain details of the original objects and operators, . . ." Thus, planning for Newell and Simon is what I have called planning by abstraction. I have chosen not to require abstracting as a criterion for planning for two reasons. I wanted to provide a slightly more general characterization of planning to include cases such as the chess example above for which abstraction does not seem to be a very important part of the planning process. More important, though, I wanted to emphasize that there are factors other than abstracting which contribute importantly to the effectiveness of planning. In particular, I wanted to emphasize the factors of economy, and flexibility discussed above.

Here is a practical example of planning by abstraction. An architect planning a hotel will typically begin with very crude drawings which take into account only the most abstract properties of the structure to be built. He or she may draw circles to indicate the general positions of the major units, e.g., the registration area, dining areas, kitchen, guests rooms, recreation areas, etc., with arrows indicating traffic flow. These drawings provide no hint either of the shape or the appearance of the structures to be built. Later in the design process, drawings become progressively more detailed and specific until the final drawings become, literally, blueprints for construction.

Perhaps the best known application of planning by abstraction in the A. I. literature is the work of Sacerdoti (1974). Sacerdoti was concerned with providing a planning procedure for a robot that had the task of moving objects from place to place in a suite of rooms. To move an object from one place to another, the robot first had to plan a path so that it could reach the object to be moved and another path from the object to the object's destination. This involved finding a path through adjacent rooms connecting the robots initial location, the object's initial location, and the object's destination, determining if there are doorways connecting the rooms, and dealing with any closed doors and furniture which may block the way. To solve this problem by abstraction, the first step would be to simplify the task. One way to do this would be to concentrate on the problem of identifying the sequence of rooms and to forget, for the moment, about the problems of doorways, closed doors, and inconveniently placed furniture. Once this simplified problem has been solved and a promising path has been identified, the path can be used as a plan to guide the solution of the original problem. That is, one can follow the suggested path, checking along the way for available doorways, closed doors, and furniture barriers.

The effectiveness of planning by abstraction comes from its ability to reduce the amount of search required to find a solution to the original problem. It does so through applying the following simple, common sense heuristic: "If an alternative doesn't meet the most important criteria for success, then it probably isn't worth checking how it does on the less important criteria." Thus, if a room isn't connected to the one where the goal is, it isn't worth checking to see whether or not its door is locked.

An important feature of planning by abstraction as Sacerdoti has described it is that tasks are simplified by dropping the less critical features and retaining the more critical ones. Thus, planning by abstraction tends to be "top down" planning, that is, planning which is shaped primarily by the top-level goals of the task. In the case of Sacerdoti's ABSTRIPS program, planning proceeds from the most important goal to the next most important goal on down to the least important goal.

Planning by analogy

In some cases, the act of representing one problem reminds us of another similar problem that has already been solved. The solution of the second problem may then prove useful as a plan for solving the first without the need of any further problem solving activity in the planning space.

Kohler (1940) reports a very interesting study of planning by analogy which shows that it is very sensitive to factors influencing the planners attention. Kohler asked people to solve an equation which involved multiplying 21 X 19. When the participants reached their conclusion, Kohler pointed out to them, apparently as an aside, that 21 X 19 is the same as (20 + 1)(20 - 1) and that this in turn is the same as 400 - 1. Later in the experiment, Kohler asked the participants to to solve an equation which involved multiplying 32 X 28. This product could be calculated, by analogy to the first problem, as (30 + 2)(30 - 2) which equals 900 - 4. In one condition, the participants solved visual puzzles in the interval between solving these two algebra puzzles. In the second condition, the participants solved other algebra problems in the interval. Kohler found that most people in the first condition (73%) solved the problem using the analogy but that relatively few people in the second condition (26%) did so. The participants who failed to solve by analogy had not forgotten the trick they had been shown. They all remembered it when asked. They simply had not thought spontaneously to apply it.

Kohler's experiment demonstrates both that analogy can be an important planning method and that the discovery of an analogy depends critically on the immediate circumstances which influence the planners attention.

Planning by modeling

Planning by abstraction gains its power by representing the original task in simplified form in the planning space. In contrast, planning by modeling gains its power by representing the original task inexpensively in the planning space without necessarily reducing its complexity. For example, low speed aircraft wings are sometimes designed by examining the performance of small scale wing models in wind tunnels. This method is useful because wing design depends relatively little on scale--a good shape for a wing is good both for small wings and big ones--but cost depends critically on scale. It is much less expensive to build a small wing than a big one.

Another example of planning by modeling is the practice of using small scale models to evaluate the appearance of an architectural project. A small scale model of a building viewed at five feet can covey useful information about what the finished building will look like from 1000 feet but at much less cost. Proportion depend relatively little on scale but, again, cost depends critically on scale. Finally, planning by modeling is common in games such as chess. When chess players plan moves mentally, they try to represent the game in all its complexity. To leave out a piece or to fail to consider a

possible move could be disastrous. The advantage of planning moves mentally is that it is less costly. A move on the board involves a commitment of resources. It can't be taken back. However, a mental move can be retracted just as an unsuccessful model of a building can be discarded with minimal waste of resources.

Multiple Planning Spaces, Layering, and Metaplanning

By the definition of planning that we have adopted above, when planning occurs, there must be at least two distinct task environments or spaces -- one for planning and one for action. However, there is no reason to limit planning to a single space. In fact, a number of artificial intelligence programs make effective use of multiple planning spaces. MOLGEN (Stefik, 1981a,b) and MACHINIST (C. Hayes, 1987) each have two planing spaces and ABSTRIPS (Sacerdoti, 1974) allows for an unlimited number of planning spaces.

MOLGEN is a program for designing genetics experiments which has three layers. The lowest layer is the action or laboratory space. The operators in this space are actions to be taken by a laboratory technician such as killing the unwanted bacteria in a culture. The next layer above, the design space, is the first planning space. The operators in this space are actions to be taken by an experiment designer such as testing a prediction or searching for an unusual genetic feature. The top layer is the strategy space. The operators in this space concern the choice of general problem solving strategies such as whether to wait for more information or make an informed guess. The strategy space contains no knowledge of genetics -- only knowledge of general problem solving strategies. The spaces in MOLGEN are layered in the sense that each layer has a major impact on what the layer below it does. Thus, the strategy layer provides a plan for how the design layer will construct a plan. Stefik calls this planning to plan "metaplanning".

ABSTRIPS (Sacerdoti, 1974) was designed to plan paths for a robot moving objects around in a suite of rooms. The program allows for planning in an unspecified number of planning spaces arranged in a hierarchy of levels of abstraction. The top layers are the most abstract and take into account only a few features judged to be the most important ones, e.g., whether two rooms are adjacent to each other. Lower levels are less abstract in the sense that in addition to the features considered by the higher levels, they also consider less important, e.g., more easily modifiable, feature such as whether doors between rooms are open or closed.

Multiple planning spaces need not be layered. For example, MACHINIST (C. Hayes, 1987) has two planning spaces which are not layered with respect to each other. MACHINIST creates plans for machining blocks of metal in much the same way human machinists do. One of MACHINIST's planning spaces is concerned with squaring, that is, with choosing three sides of the block to smooth. The other space is concerned with choosing the order in which cuts should be made in the block. Plans are made independently in these two spaces (They can be made in either order.) and then combined into a final plan for machining the block. We will call planning spaces such a those used in MACHINIST as parallel rather than layered.

Interleaving plans and action

The planning and execution of a task need not occur in completely separate and sequential phases. It is possible and, as we will see, often desirable to plan a little and then execute a little and then plan a little more and so on. McDermott (1978) has called this "Interleaving of plan and action.

Features of the Task and Task Environment which Influence Planning

In this section, we will discuss planning in three situations that can radically influence the way planning is carried out: planning in a changing environment; planning under uncertainty; planning for decomposable tasks; and planning when costs in the action space are very high.

Planning in a changing environment

To this point, we have been talking about planning as if it occurred in a static world. In some cases, of course, the world does hold still long enough so that a planner can treat it as static. However, as Hayes-Roth and Hayes-Roth (1979) have pointed out, aspects of the world important to our task often change <u>while</u> we are planning. To understand planning that is carried out in a changing environment, we must understand the sort of inputs from the environment that may influence planning and the sorts of tasks in which the planning process is most subject to these changing inputs.

We will discuss four sources of input during planning: Changes in the external environment, changes due to the planning activities themselves, and changes due to the execution of the plan.

Changes in the external environment

Planning carried out in industrial settings is frequently subject to unpredictable changes initiated by management and by market forces. Planners of hardware or software or instructional manuals may be told in the midst of their activities that management has decided to add a new feature to the product or to drop it altogether. Unpredictable inputs such as these are important but we can do little about them. Of more interest to us are the sorts of external inputs from clients and collaborators which can be anticipated to some extent within the planning process. For example, architectural and industrial design tasks typically involve a client who periodically evaluates the results of the design process. The designer-client relationship, a kind of cooperative game, shapes the planning process by leading the designer to plan only up to those decision points that are the responsibility of the client. Thus, the designer tends to put little planning effort into alternatives which the client may reject, saving that effort until after the client has indicated that the alternative is worth developing. The impact of the designer-client relation is important in the case study to be discussed below.

New inputs stimulated by planning activities

As noted above, the activity of planning may result both in further specification of the goal and in the construction of a plan for accomplishing the goal. Tasks which require the planner to do a great deal of goal specification have been called "ill-defined tasks" by Reitman (1965) Ill-defined tasks are very common in a variety of fields including writing, architecture, and software design. The activity of specifying the goal which is required in ill-defined tasks can result in important changes in the planning process. For example, a client may ask an architect to design a modern office building which fits on a specified lot and provides a specified amount of office space for a particular business. This is clearly an ill-defined task because to create such a design, the architect has to make a very large number of decisions such as specifying the organization of the space and the placement of stairs and elevators as well as the treatment of the windows and the style of the lobby. Planning in tasks such as these is very likely to be influenced by the decisions which the architect makes during planning. Thus, early decisions about the placement of the stairs may have to be changed because they create difficulties for the design of the lobby. These changes may in turn require changes in the organization of the office space, etc.

A second sort of input from the planning process is generated through the evaluation of plans. As Hayes, Flower, Schriver, Stratman, and Carey (1987) point out, a writer may create a plan and then evaluate and reject it before any attempt has been made to execute it. For example, Kaufer, Hayes, and Flower (1986) observed a writer who proposed early in planning to present a sequence of examples and then, long before anything had been written, changed the plan on the grounds that what he had proposed would be boring.

New inputs stimulated by attempts to execute the plan

The attempt to execute the plan may reveal faults that were not evident at the time the plan was conceived. For example, Kaufer et al. (1986) described a writer who planned to write a section on the history of a topic and a following section on the "current status" of the problem. After writing the history section, however, he discovered that what he had to say about current status was redundant with what he had just written. As a result, he dropped the "current status" section from his plan.

There are two sorts of tasks in which it appears especially likely that attempts to execute the plan will stimulate new inputs to the planning process. These are construction tasks and resource limited tasks.

Construction Tasks. By construction tasks, I mean tasks such as building a porch, painting a picture, and writing an essay. What these tasks have in common is that each creates a tangible product -- a product that is produced through the cumulative effect of a sequence of actions. Thus, the product takes shape as the task proceeds and may change the task environment in very important ways.

For example, writers frequently consult the text that they have just written for ideas about how to proceed. Kaufer, Hayes, and Flower (1986) have shown that in composing sentences, writers frequently reread the beginning of an incomplete sentence in order to get a 'running start' in composing the next segment. In design, the designer's current sketches are used as a source for ideas and inferences which shape later design decisions. For example, in one design episode, Ballay et al (1984) observed a designer examining a crude model of a proposed solution and discovering that it would be easier for right handed people to use it than left handed people. The designer thus was alerted to the issue of handedness which then became an important criteria by which to judge later solutions.

Because the partially completed product is part of the task environment and because it is continually changing, the task environment is continually changing in construction tasks. These changes in environment may provide new inputs to the planning process in the form of new ideas about what to do and how to do it. As a result, invention may be stimulated continuously throughout the course of the execution of such tasks. Indeed, it has been observed both in writing and design tasks that invention occurs continually in these tasks right up to the moment when the final drawing or final draft is being completed. (See Hayes et al.,1987, and Ballay et al., 1984).

Information overload tasks. There are many cases in which we have to perform a task before we have learned all of the relevant task information. This may come about either because the task is very complex or because we have limited learning time. I will call such tasks "information overload tasks". A clear example of an information overload task is Hayes-Roth and Hayes-Roth's (1979) errand planning task. Participants in this task were given the following instructions about tasks to be completed in an imaginary town:

> 'You have just finished working out at the health club. It is 11:00 and you can plan the rest of your day as you like. However, you must pick up your car from the Maple Street parking garage by 5:30 and then head home. You'd also like to see a movie today, if possible. Show times at both movie theaters are 1:00, 3:00, and 5:00. Both movies are on your "must see" list, but go to whichever one most conveniently fits into your plan. Your other errands are as follows:
>
> o Pick up medicine for your dog at the vet.
> o Buy a fan belt for your refrigerator at the appliance store.
> o Check out two of the three luxury apartments.
> o Meet a friend for lunch at one of the restaurants.
> o Buy a toy for your dog at the pet store.
> o Pick up your watch at the watch repair.
> o Special-order a book at the bookstore.
> o Buy fresh vegetables at the grocery.
> o Buy a gardening magazine at the newsstand.
> o Go to the florist to send flowers to a friend in the hospital."

The participants were also given a street map of the imaginary town, measuring roughly four blocks by six, on which were located about 90 places of potential interest. The authors do not report how much time participants were given to examine the map. However, from the sample protocol they provide, it is clear that the one participant whose performance is described in detail did not learn all of the relevant task information that was available in the map before starting the task.

Hayes-Roth and Hayes-Roth claim that planning in the errand planning task is "opportunistic". They describe opportunistic planning as planning that is driven by events encountered by the participant while engaging in the task. They draw a sharp contrast between opportunistic planning, which they characterize as bottom-up planning, and planning by abstraction, which they characterize as top-down planning. They provide a complex model of opportunistic planning based on the architecture of the HEARSAY program (Reddy & Newell, 1974) and use it to account for the behaviors observed in the sample protocol.

Although Hayes-Roth and Hayes-Roth's planning model has attractive features, I believe that it does not provide a thoroughly satisfactory account of planning activities. I propose the following as a more reasonable account of performance in the errand planning task, and more generally in information overload tasks.

Participants examined the map and identified their starting place, the health club, their final destination, the Maple Street parking garage, and some locations where the most important errands could be accomplished, e.g., vet's offices and appliance stores. Thus, the participants might be described as creating a crude mental image of important map locations and using that image to plan moves. Clearly, such activity could be characterized as planning by abstraction. Next, the participants used their plans to guide action. In this task, taking action means tracing a path visually on the physical map between locations for planned action, e.g., between a veterinary hospital and an appliance store. As participants trace these paths, they may discover unexpected opportunities to accomplish secondary goals, e.g., to pick up a gardening magazine at a news stand discovered along the way. These opportunities appear unexpectedly <u>because</u> this is an information overload task in which participants do not have the time or resources to learn all of the relevant task information. Further, the

nature of the task may leave the participant little choice of planning strategy other than planning by abstraction, that is, planning based on a simplified image of the available information.

This characterization of performance in the errand planning task is at variance in some important ways with the characterization provided by Hayes-Roth and Hayes-Roth. These authors drew a sharp contrast between planning by abstraction which they described as "top down" and planning in the errand planning task which they described as "opportunistic" and "bottom up". In the characterization of performance in the errand planning task that I have presented here, the primary planning activity that participants engaged in was planning by abstraction. This planning was focussed on accomplishing the most important of the errands. In addition, the participants discovered information during attempts to execute a part-plan in the action space (planning and action were interleaved) which allowed them to accomplish some less important errands as well.

There seems to be some confusion in the Hayes-Roth and Hayes-Roth monograph in the use of the t erms top-down and bottom-up. Planning by abstraction was characterized as top-down because it focussed on the most important goals first and less important ones later. However, opportunistic planning was characterized as bottom-up not because it dealt with the least important goals first but rather because it was data driven. While it is not uncommon to view data as low level detail and hence to see it as not "important", it is not appropriate to apply this view to the errand planning task. In that task, data discovered while executing a plan could equally well suggest ways to accomplish important tasks as unimportant ones.

Tasks involving risk and uncertainty

The cost effectiveness of planning many steps ahead may be reduced if the outcome of the steps is uncertain. In such cases, more interleaving of plan and action may be appropriate. If the planner believes the probability that planned actions will have their expected outcomes is only 0.8, then a sequence of four such actions would have a probability of success of only 0.4 and a sequence of ten actions, only 0.1. In such uncertainty, the planner may be unwilling to plan long sequences of action without testing them. Rather. the planner will prefer to plan short sequences and execute them to be sure that the plan is on the right path. Thus, planning and action may be interleaved because the planner is uncertain of the effect of actions being planned.

For example, imagine an architect who is designing two buildings, one for each of two clients she has known for years. Client A has almost always liked and accepted the architects' suggestions but client B, over the years, has rejected about half of the architects suggestions. For client A, the architect identifies a potential site, draws a site plan for the building, and draws several sketches of the proposed building for discussion at their next meeting. For client B, however, all she does is to identify a potential site because she feels that, due to the unpredictability of the situation, the planning effort involved in drawing the site plan and the sketches is liable to be wasted.

The relation of cost and layering

Tasks which involve large expenses such as constructing a skyscraper or manufacturing an automobile, may warrant planning activities which are themselves expensive, e.g., the construction of very precise models or the production of numerous detailed high quality drawings. The costs involved is these planning activities may themselves warrant planning and so on. Thus, the larger the commitment of resources that the original task involves, the more likely it is that planning for the task will be multi-layered.

We expect that early in the process of solution, problem solvers will work toward solution using relatively inexpensive representations -- rough sketches, notes, outlines,etc., and later, work work with progressively more expensive representations, e.g., more precise drawings, drafts, etc., before committing themselves to a solution in the most expensive representation, the final external product. This progression from crude, inexpensive representations to more precise and expensive ones is an important consequence of the nature of tasks.

Summary

Our analysis of the research literature on planning has allowed us to synthesize a useful theoretical perspective on planning activities. This perspective helps us to clarify such questions as "What processes should be subsumed under planning?", "Why is planning useful?", "What forms does planning take?", and "How do features of tasks influence the way planning is carried out?" In research now under way, we are applying this theoretical perspective in an attempt to understand the similarities and differences between planning activities as they are carried out in industrial design and in

writing tasks. Our long term objective is to understand how task variables in writing and other complex tasks influences performance in those tasks.

Section 2: Observational Studies of Planning in Writing

Differences in Writer's Initial Task Representations

The purpose of this exploratory study was to search for relations between the quality and quantity of planning carried out prior to writing and the quality of the written product.

Subjects. The subjects of this study were 12 writers with varying degrees of experience. Five were expert writing teachers and seven were college student writers.

Task. The subjects were asked to write about their job for Seventeen Magazine. They were told that the readers of the magazine would be teenage girls (aged thirteen to fourteen). The subjects were given an hour to complete their tasks, and verbal protocols were recorded as they worked.

Method of analysis. The texts were rated for quality by four experienced writers who considered the following three dimensions of the essays:

How well is the text adapted for the audience?

To what extent does this text have a clear point, focus, or rhetorical purpose that goes beyond simply "knowledge telling" on a topic?

How well constructed is the text in terms of overall organization and coherence?

Analysis of planning was limited to the initial portion of the protocol which ended when the subjects wrote their first complete sentence. The protocols were divided into clause units which were coded into the following categories: Reading/paraphrasing the task instructions, process goals, metacomments, and planning. The analysis focused on those statements coded as planning. [In light of our discussion above, in which we include goal setting as an integral part of planning, it might have been preferable in this analysis to include process goals as part of planning.]

The planning clauses which were not repetitions of earlier clauses were counted. In addition, the planning clauses were rated for quality by two experienced raters on the following three dimensions:

How well does the writers planning reflect a concern for audience?

To what extent does the writer's planning reflect a concern with developing a clear point?

How far does the writer's planning reflect a concern with structuring the text or fitting a genre?

Results. Surprisingly, no significant relation was found between the writer's experience and either the amount or the quality of the initial planning revealed in the protocols. Nor was there a significant relation between the writer's experience and the quality of the texts they produced. However, strong relations were found between the quantity and quality of initial planning and the quality of the resulting texts. The quality of the texts correlated 0.655 with the quantity of planning and 0.874 with the quality of planning.

These results are exploratory and their appropriate interpretation is not yet clear. At one extreme, one might want to take these results as indicating that we should teach writers to do more and better planning so that they will write better. On the other hand, the relations among these variables may be mediated by some third variable such as motivation. Some writers may have taken the task more seriously than others and therefore worked harder at all aspects of the writing task. This second interpretation is rendered more plausible by the lack of relation between experience and performance.

Planning in Writing: The Cognition of a Constructive Process.

This speculative article proposes several characterizations of the processes which adult writers bring to ill-defined expository writing tasks. First, it proposes that there are three executive level strategies which writers may bring to planning: knowledge-driven planning, script- or schema-driven planning, and constructive planning. In knowledge-driven planning, the writer can use the current organization of information in memory as a plan for sequencing topics. In schema-driven planning, useful with familiar genre such as children's stories, the writer can draw on a preformed schema to help in the selection of goals, tests, and plans. In constructive planning, writers set their own goals, criteria, plans, and procedures in response to the task. In addition, the article presents a characterization of constructive planning based on analysis of expert and novice writers. It isolates five critical features of the constructive planning strategy in which writers must create a unique network of working goals and deal with the special problems of integration, conflict resolution and instantiation which this constructive process entails.

Section 3: Skills Involved in Planning to Create and Planning to Revise Texts

In planning to create a text, writers must make informed choices among writing strategies. For example, they must when an the insertion of an example, a metaphor, or a heading will be helpful to the reader. In planning to revise a text, writers must be able to evaluate whether or not a text will have the intended effect on the audience. For example, they must be able to perceive that part of a text contains ambiguities or fails to provide the reader with needed information.

The two studies described in this section explored these issues. The metaphor study explored writers' abilities to predict the effect of metaphor on an audience. The knowledge effect study demonstrated that providing topic knowledge to writers prevented them from perceiving that a text failed to provide that topic knowledge to the intended readers.

Designing Instructional Texts: Metaphor as Explanatory Strategy

Science writers, and writers of instructional texts of all sorts, face a doubly difficult task in introducing readers to complex and unfamiliar topics. First, they must correctly identify those parts of the material that their readers will have difficulty understanding. As Swaney, Janik, Bond, & Hayes, (1980), Schriver (1987), and Hayes et al. (1987) have shown, writers at all skill levels have difficulty in predicting which aspects of a text will be difficult for members of the intended audience. Second, when they have identified a part of the material that will confuse the readers, writers have to choose an appropriate procedure for clarifying that portion of the text. Clarifying procedures which are frequently used include examples, graphics, and metaphors. Leatherdale (1974) has pointed out that science literature makes extensive use of metaphor. Extensive use, however, is not evidence that the choice of metaphor is always the right one to make, or that, as applied, metaphor is always effective in promoting readers' comprehension of the topic. In this paper, we will explore the accuracy of writers' decisions in electing or rejecting the use of metaphor and the effectiveness of two types of metaphor in promoting readers' initial learning of science topics.

Metaphor, simile, and analogy

ideas or objects that are usually unrelated (MacCormac, 1985). For example, in these lines:

> That time of year you mayst in me behold,
> When yellow leaves, or none, or few do hang
> Upon those boughs which shake against the cold,
> Bare ruin'd choirs where late the sweet birds sang.

Shakespeare (Sonnet LXXIII) compares the appearance of an aged man to that of a tree in winter and compares a branch used as a perch by birds to a choir loft.

Closely related to metaphor are simile and analogy, terms which also suggest comparisons. A simile is a metaphor which includes one or both of the comparative terms "as" and "like" as in the opening lines of Shakespeare's sonnet LX:

> Like as the waves make toward the pebbled shore,
> So do our minutes hasten to their end,

Both metaphors and similes convey a very wide variety of comparisons. These include the comparison of appearances, e.g., sparseness of leaves and thinness of hair, shaking of tree limbs and palsied movement, as well as the comparison of complex relationships, e.g., the motion of a wave toward the shore and the passage of life toward death.

There is some disagreement among theorists about the relation between analogy and metaphor. For MacCormac (1985), analogies include comparisons between things that are similar as well as things that are not. He sees the comparison of a ship to a model of that ship as an analogy but not as a metaphor because the comparison is not a surprising one. Thus, for MacCormac, metaphor is a subset of analogy. For Gentner (1980), however, the reverse is true. In her view, analogies are always based on relations and not on appearances. For example, although water and electricity don't look alike, one can serve as an analogy for the other because they have comparable relations, e.g., water pressure is a measure of water current as voltage is a measure of electric current. In contrast, metaphors include comparisons based on attributes such as color, size, appearance, etc. as well as comparisons based on relations. Comparing a red-cheeked face to an apple would be a metaphor for

Gentner but not an analogy. Thus, for Gentner, analogy is a subset of metaphor.

Although it may be useful to distinguish between analogy and metaphor for some purposes, we don't believe that the distinction is important for the task of writing instructional texts. If writers call readers' attention to a comparison between ideas that they regard as similar, that very act suggests that they believe the similarity is not obvious to the readers. Thus, from the point of view of the reader, all comparisons made for instructional purposes are non-obvious and therefore metaphoric by MacCormac's definition. Further, though a distinction can be drawn between comparisons of attributes and comparisons of relations, we don't believe that the distinction is an important one for writers when they are choosing explanatory strategies. In this paper, therefore, we will not distinguish among analogy, simile, and metaphor because we believe that they function in essentially the same way in instructional texts.

Metaphors: dead and alive

The purpose of metaphor in instructional texts is to help readers learn about a new topic by comparing it to one they already know about. The new topic is called the "target" and the already known topic, the "base." (In literary terms, these are the "tenor" and the "vehicle.") Rutherford's model of the hydrogen atom is a metaphor which explains the structure of the atom (the target) by comparing it to the structure of the solar system (the base). In this model, electrons are pictured as small objects circling a much heavier atomic nucleus just as planets circle a massive sun.

Although we have described the use of metaphor as a conscious strategy employed by writers who are trying to solve problems of clarity, Lakoff and Johnson (1980) argue that metaphor is a frequently unconscious and extremely common feature of everyday language and that we use it routinely often without being aware that we are doing so. They note that many common objects cannot be described without metaphor, e.g., we speak of the leg of a table and the hands of a clock, and, further, that some verbs act metaphorically by personifying inanimate objects and projecting intention and agency upon them. Consider their example: "Inflation is lowering our standard of living " (26).

In contrast to Lakoff and Johnson's (1980) position, MacCormac (1985) considers that only those comparisons which create semantic anomaly, i.e., comparisons perceived as strange or unusual, are properly called metaphor. He excludes language that is taken "under normal circumstances in a discourse community to be a normal utterance with no semantic anomaly" (130). Similarly, Leatherdale (1974) holds that metaphor creates a disharmony between words and content that is, at first sight "incongruous in some respects" and that this dissonance motivates a kind of closure-seeking, motivates the reader to search for similarities between the two domains, and thus promotes learning (98). As Kittay (1987) puts it, the function of Metaphor is ". . .to provide a perspective from which to give an understanding of that which is metaphorically portrayed." (3,4) She explains that one component of metaphor (what we would call the base) can be used to conceptualize the other (the target). (25). These views are compatible with the definition given earlier, that is, that a metaphor is a comparison between two things that are ordinarily considered unlike and that a metaphor's function is to point out similarities not previously considered.

We should note that Lakoff and Johnson are concerned with phenomena that are quite different from the ones we have focused on in this paper. Lakoff and Johnson are concerned with the largely unintentional comparisons which can (rightly or wrongly) be inferred from a literal reading of everyday language. In contrast, we are concerned with writers' intentional use of comparisons which they believe their readers would be unlikely to think of on their own. While Lakoff and Johnson have expanded the definition of metaphor to include dead metaphors as well as explicit and obvious comparisons, we choose to work with MacCormac's more restricted definition.

Ambiguity in metaphor

The readers' response to a metaphor may be hard to predict if the base or the target or both are complex. For example, there are many properties of old men and trees which might be compared to each other, e.g.,sparse leaves and thinning hair, bark and skin, roots and feet, upright posture, etc. If the writer were concerned with emphasizing a particular comparison, the availability of alternative comparisons could be a problem. In poetic texts, however, this multiplicity of meanings is valued. Metaphor is often used liberally to explode meaning. The ambiguity of metaphor invites the reader

to associate freely to a variety of experiences. Indeed, an accepted characteristic of literary texts is that they provide fertile ground for calling up personal experiences and constructing a plurality of possible meanings. The readers' free associations and idiosyncratic inferences, facilitated by the ambiguity of metaphor, are welcomed.

When encountering metaphor in a literary text, the reader is assumed to be familiar with both the base and the target, e.g., we know about both old men and trees, about the passage of time and the progress of waves. Readers are not expected to treat the target as a new idea to be learned (for they are already familiar with it in many ways) but rather to become aware of parallels between base and target that they had not previously noticed.

Metaphor is used quite differently in instructional texts. Here, the reader is expected to know the base but not the target. The writer's intention is to direct the readers' attention to some properties of the base but not to all of them. In instructional text, metaphor is intended to spotlight particular properties of the base rather than to floodlight them all. In this sense, the function of metaphor is not to explode meaning but to constrain it by focusing a readers' attention on the target information and discouraging readers from attending to associations that may be erroneous or distracting.

Single and multiple metaphors

There is fairly wide agreement in the literature that metaphor is a useful explanatory device (Gentner, 1980, 1982; Gick and Holyoak, 1980). However, there is mixed advice on the best practical course to follow in using metaphor in instructional texts. Some experts point out the advantages of covering the topic with a single, comprehensive metaphor, while others suggest covering the topic with many limited metaphors, each illuminating one or a very few features of the topic. Gentner (1980), using examples of historically important metaphors, prefers single, comprehensive metaphors, in which a number of features from a base system can be accurately mapped onto corresponding features in the target material. Her examples suggest that writers would do well to choose a large, comprehensive base that can achieve systematic coherence with a target domain. Carroll and Thomas (1980) also recognize this advantage, noting that, "The more aspects of the target system that can be 'covered' by a single metaphor, the better" (7).

In contrast to a single metaphor approach, Spiro et al. (1987) advocate the use of multiple metaphors. They report that commonly used metaphors in medical education settings have both positive and negative effects. That is, although metaphors used to describe physiological processes do promote correct understanding of some features, they also lead to misunderstanding in several ways. First, they may lead students to oversimplify a complex target domain. That is, students may remember only those properties of the target domain that are covered by the metaphor and may ignore important features not covered by the metaphor. For example, medical students may think of the circulatory system as a plumbing system. The fact that plumbing systems typically have rigid pipes leads students to an oversimplified view of blood flow which ignores effects due to the elastic nature of blood vessels. Second, students may attribute more of the features of the base to the target than the writer intended. For example, by making an inappropriate analogy between heart muscle and skeletal muscle, medical students often misinterpret the enlargement of the heart in congestive heart failure as resulting from the stretching of heart muscle beyond its elastic limit by overwork -- a mistaken view. Spiro et al. (1987) list eight different ways in which metaphors may lead to misunderstanding. Their solution to the problem of metaphor induced misunderstanding is to suggest that writers should use multiple metaphors, each specialized to convey particular features of the target and, where possible, designed to correct misunderstandings conveyed by other metaphors. Carrol and Thomas also note this drawback of the single metaphor approach, recommending that writers wean their audience from extended metaphors when they no longer apply and either revert to literal discourse or switch to a new, more accurate metaphor.

The difference between these two views does not appear to be a difference of opinion about whether it would be useful for instructional purposes to have a metaphor which conveyed all of the to-be-learned features of the target and no irrelevant features. Rather the difference appears to be a difference of opinion about the probability of finding such metaphors. Gentner, working from some elegant historical examples, e.g., Rutherford's model of the atom, takes an optimistic view, and Spiro et al. working in a practical medical school setting, take a pessimistic view. Carroll and Thomas (1980) take an intermediate view, recommending the use of single, inclusive metaphors generally but recognizing that there are situations in which multiple metaphors are needed.

Although the literature offers us many hypotheses about how single and multiple metaphors function, little is known about the effects of metaphor in actual reading-to-learn tasks. Carroll and Thomas aptly dub their suggestions "best-guess recommendations" in that their recommendations emerge from cognitive theory, but have not been systematically tested on actual readers. While Gentner's criteria for good analogy come from examples that have proven successful over the years, we have yet to see readers' reactions to texts written according to criteria inferred from these historical examples. And although Spiro et al. tested metaphor on actual readers, they observed effects on reader's advanced learning rather than the initial learning of concepts.

Clearly, the literature has provided mixed reviews of the costs and benefits of these explanatory strategies, but what do actual writers believe and how do actual readers respond to metaphor? User testing may be a fruitful way of further exploring and evaluating the suggestions of these researchers. In this paper, we will present three sorts of observations bearing on the instructional use of a single, broad metaphor and of multiple, narrow metaphors: First, we will present a study of writer's predictions about the relative instructional effectiveness of science texts written without metaphor, with a single metaphor, and with multiple metaphors. Second, we will describe a study comparing student learning from these three kinds of science texts. Finally, we will discuss our experience as writers designing texts with and without metaphor.

Study one: Writers' predictions of readers' response to metaphor in technical texts

In this first study we asked technical writers at three different levels of experience to evaluate the effectivness of three introductory science topics, each written with various kinds of metaphor. Our purpose was to observe and compare these writer's predictions about the efficacy of metaphor in technical texts.

Subjects. Our novice writers included nine seniors and eight master's students majoring in technical and professional writing at Carnegie Mellon. Our experienced group was comprised of nine technical writers with at least five years of experience in the field.

Materials. The materials packet contained nine instructional texts covering three science topics: electricity, pattern recognition theory and the immune system. We prepared or borrowed short, two to four page "metaphor free" versions for each topic (we called this version P, for "plain") and then revised each plain text to produce two additional versions of each topic-- one employing several ad hoc metaphors (we called this the AH or "ad hoc" version) and one using a single, extended metaphor (the E or "extended" version). Figure 2 summarizes our design of these materials. The nine texts are included in Appendix A.

All texts were roughly two to four pages in length; however each topic presented a different number of target ideas. The immune system texts were the most dense, each presenting a core of 34 target ideas. The perception texts each covered 13 target ideas, while each electricity text covered seven. The extended version of the immunity text employed a war metaphor, while the extended electricity text discussed current in terms of crowd behavior. The extended, perception text employed Selfridge's (1959) "pandemonium" metaphor. These extended metaphors were applied throughout each topic. In the ad hoc versions, we opportunistically created metaphors suggested by the local context of concepts in the texts. For example, in the immunity text, we compared the basophil to a smoke detector that signals a fire alarm whenever a pathogen is detected in the body. This smoke detector metaphor was used just to explain this concept, and was not connected to other metaphors used to explain other concepts in the text. In the AH versions, roughly half of the target concepts were covered with short, ad hoc metaphors. We will discuss the design of these texts further later in this paper. A rank and comment sheet was attached to each packet of texts.

Procedures. Each writer read all nine texts which were separated into groups according to topic. The order of the passages within each set were counterbalanced according to version (P, AH, E); the sets themselves were counterbalanced according to topic. The writers were asked to rate topic for difficulty and then to rank each version of the topic from one (for low) to three (for high) to indicate its suitability for the audience fo first year college students with little prior knowledge about the topics. Writers were asked to imagine how well students would learn and recall the information in these versions.

In assessing difficulty of topic, these writers rated the electricity topic as easiest, followed by the pattern recognition topic and then the immunity text, which was considered the most difficult.
Figure 3 shows the average of the writers' rankings of the various text versions. While differences among the various versions are not statistically significant, there appears to be a tendence for the writers to judge the metaphoric texts as more appropriate for the audience than the non-metaphoric ones.

Study two: Effects of metaphor on learning

Thus far we have considered researchers' mixed advice on the use of metaphor and writers' predictions about the effects of metaphor on readers. Although both groups view metaphor optimistically, i.e., as a helpful device in writing, both writers and researchers have mixed opinions about which kinds of metaphor will be most effective. This study tests students' ability to learn from the three kinds of metaphor in the science texts designed for study one. As Schriver et al. (1986) have noted, this kind of reader-based testing can help writers make more informed choices about clarifying devices such as metaphor, and can help writers better predict the effects those choices will have on real readers.

Subjects. Subjects were nine males and nine females, all first year Arts or Humanities majors at Carnegie Mellon. They received course credit for their participation. Science and technical majors (e.g. engineering) and students taking biology or physics were instructed not to sign up.

Materials. Materials included the science texts used in study one. These texts were placed into packets for each of the students, so that each packet included three texts, one covering each topic. One of these texts was in the extended condition, one in the plain, and one in the ad hoc condition. The order of these texts was counterbalanced across subjects.

The recall interview questions constructed for each topic included an initial, general question (e.g. In general, what did this text teach you about; that is what were the main points this texts tried to convey?) followed by several more specific questions on the target ideas in each text (e.g., Describe all the different ways that pathogens can harm your body). Participants were tested for recall

of a set of core relations and operations described in the texts. For example, here are some target ideas from the immunity text:

- protozoa attack red blood cells
- agglutin clumps pathogens together
- vaccination is the injection of a virus

Procedures

In an initial session, students were given the packets and told to read the contents carefully, as they would be asked to recall and explain the materials the next day. Students were instructed to focus on the key operations in the texts and were informed that the student remembering the most information would be awarded a $10 prize. Students had 30 minutes to read the three texts. The experimenter gave a signal when half of the time remained, and again when one minute remained. Students were told they could reread, study and underline concepts, but could not take notes or bring the texts with them the following day.

In a second, 30 minute session the following day, the experimenter interviewed each student individually, proceeding from general to specific questions. Students were told to say as much as they could remember. At various points, the interviewer asked for clarification and prompted students for further information, until all the questions were exhausted and students remembered, misremembered or said they could not recall anything more. These interviews were tape recorded.

Analysis. The present study is designed to tell us which type of metaphor (if either) may be more or less effective in promoting recall of target information. We scored our subjects' recorded explanations to determine the percentage of target ideas remembered from each of the three conditions. Each correct answer received one point. Mistakes or errors were also tallied. This enabled us to compare total recall scores across the three versions of each text. Scoring for error also allowed us to locate and examine sources of error.

Because we noted that a high percentage of students were recalling certain common sense concepts that even non-technical majors may have already understood (e.g., vaccination is an injection of a virus), we conducted a second analysis scoring only for the "hard" concepts,

those concepts which were correctly remembered by fewer than 72% of the students.

In addition, we reviewed the recall tapes to determine whether the metaphors reappeared in students' explanations.

Results. In general, writers' predictions about the difficulty of the three topics was quite accurate (See the right hand column of Figure 4); readers in the present study found the electricity topic the least difficult, followed by the immunity text. The perception topic was the most difficult to learn.

An analysis of student recall indicated that for all three texts, there was no significant difference between number of concepts recalled in the metaphor and non-metaphor versions, nor did the extended and ad hoc versions prove more successful in enhancing recall. Surprisingly, our manipulation of metaphor in the texts seemed to make little difference in the mean number of total propositions recalled, as indicated in Figure 4. Overall, 41.2% of information in the plain versions was recalled, 38.9% of the mixed versions and 33.6% of the extended versions.

insert table

As Figure 5 indicates, both versions with metaphor tended to produce a larger number of errors in student recall.

Discussion. Our study suggests that metaphor does not always enhance the learning of new concepts, but may lead to more error in recall. How do we understand this increase in error?

Spiro, Feltovich, Coulson, & Anderson (1987) have shown that error can occur when readers make faulty inferences about a topic, overextending the metaphoric base in inappropriate ways. However, the error produced in our study could not be exclusively linked to misuse of metaphors. In fact, we were quite surprised at the infrequency with which students referred to our metaphors as they explained concepts in the recall interviews. The type of errors our students made tended to involve confusion with target concepts (e.g., confusing the role of a basophil and opsonin, or combining the role of the first process in the perception model with that of the second).

It seems more plausible that error in our study might have been a result of simple distraction or a result of the additional processing involved in reading metaphor. In that metaphors are elaborations of key concepts, they may distract readers from information that is more central yet less interesting or vivid. Indeed, some of our metaphors were more vivid (e.g., the war analogy, the demons) and perhaps more interesting than the topic material they covered. It may also be that students had more information to process in the metaphor conditions, and that this may have taxed their ability to study the important target information and remember it correctly. Although the extended versions were not longer than the metaphor-free versions, they were denser in terms of ideas, if one considers the base information and comparisons in addition to the target ideas for which we scored. This seems especially true in the extended immunity text, the densest of all in terms of target ideas and number of references to the war base. This text was also highest in error.

But the more intriguing question in this study is why metaphor did not live up to writers' expectations. Why did it fail to enhance learning? As we have already speculated, the extra processing involved in reading metaphor may take its toll. But we might also consider that metaphor was not stable across these texts, and in this sense, our comparisons were perhaps not fair to begin with. For example, each of the "extended" versions differed in the amount of base information that could realistically be brought into the text and the way this base information could be staged and mapped onto the target information. There were some cases in which we failed to find a way in the extended versions to make the base "cover" all the target material. In other cases, e.g., the immune text, the base was more flexible and we were able to cover more concepts throughout the passage. Because of the opportunistic nature of finding and applying appropriate ad hoc and extended metaphors we could not reliably ensure that all the target concepts in all the texts received the three different kinds of treatment: no coverage with metaphor, some coverage and extended coverage. For example, consider the concept:

- the hookworm enters the body by piercing the skin.

Although this target concept was covered with metaphor in the extended version of the immunity text (hookworm is an enemy soldier that invades one's body (homeland) by aggressively piercing

the skin (borders)), a metaphor did not present itself as we wrote the ad hoc version. Therefore, the ad hoc version presented this concept exactly as it appeared in the plain version-- uncovered by metaphor. Although this appears to be a mistake in terms of the intended comparisons between these texts (i.e., the design of our research materials), this was not a mistake on our part as writers. It became an inevitable if not sensible move to use metaphor only where it seemed to fit. Had we covered all the target ideas with discreet metaphors in the ad hoc versions, the text would have appeared disconcerting and unnatural if not insulting to a mature reader. Moreover, there were instances in which metaphors simply did not come to mind. Because our decisions as writers affected the integrity of these versions and subsequently our ability to compare them, and because our experience in writing metaphors made us aware of other variables at work in this study, we will end this paper with some discussion of the design of these texts.

In his great treatise on Rhetoric, Aristotle (ref) cautions that metaphor-making is the one thing that cannot be taught. Indeed, when we attempted to use extended and ad hoc metaphor prescriptively, as specifications for writing, we found that it was sometimes impossible and nearly always difficult to follow the advice of theorists, whose models and suggestions did not always account for the choices and considerations that we, as real writers, faced. Here we discuss some of the factors that affected our decisions as we designed metaphors for this study.

Although Deidre Gentner helps us as account for the beauty and success of some established, comprehensive analogies, these extended metaphors as we have called them, are not always plausible options for writers in that they are extremely hard to find. In fact, finding apt metaphors is probably an exception rather than a rule. That may be what makes Gentner's elegant examples so unique and so memorable.

Factors involved in designing comprehensive metaphors
Gentner claims that what makes extended analogies so useful is their comprehensiveness; in her examples, new concepts in the target domain are accurately mapped onto key relations in a base or familiar system. In writing extended versions, it became our goal to find or produce comprehensive metaphors, to attain the kind of exact mapping between domains that both Gentner and Carrol and Thomas have advocated. However, because extensive, accurate metaphors

were not always available, we had to compromise. One alternative was to create an imaginary base that would fit the target material. We borrowed an imaginary base designed by Selfridge (1959), in order to explain the letter perception passage. Selfridge's imaginary world of pandemonium is comprised of shouting red demons who compete for attention as they signal the presence of visual features which they have recognized. This invented world allowed us to be comprehensive in that it adequately covered all the main processes in the perception text: recording images, discriminating among visual features, counting and signaling the presence of features and patterns. It also had visual impact. However, invented or imaginary bases such as Selfridge's can fail to meet one important criterion for effective metaphor. As Gentner points out, the metaphorical base should be familiar and predictable for the reader, for it is this stable and well-structured set of relations onto which readers will map new and unfamiliar target material. The world of shouting, signaling demons who engage in recognition tasks is not, however typical, well-structured or predictable. Many imaginary bases will clearly lack these qualities, unless the reader is already familiar with this invented world.

Another option for attempting comprehensiveness was to use a generally comprehensive and familiar base (such as the war scenario we integrated into our immunity passage); but to revert to literal language to explain target concepts that don't fit the metaphor. Although this strategy is in keeping with the advice of theorists who recommend that metaphors should dropped when they no longer apply, this strategy raised some concerns for us. In short, it made the coverage of target concepts seem rather arbitrary; in some cases key concepts were not "covered" by the metaphor, while other concepts were. We worried that, as in Spiro et al.'s study, this strategy would put unintended emphasis on some concepts, while underemphasizing others. This problem was even more apparent as we wrote the ad hoc metaphors, which were always applied opportunistically. For some crucial concepts, no quick metaphor came to mind. As a result, the number and placement of metaphors in the ad hoc version was not consistent for the three topics. It was hard to prescribe a fixed number of metaphors to cover the important concepts for each text without stretching the appropriateness of the metaphors we chose to use.

In order to investigate the recall differences between concepts covered with metaphor and concepts explained literally, we

ran an additional analysis, comparing covered concepts in the ad hoc versions with the same concepts (but untreated by metaphor) in the plain versions. Figure 6 reports the mean number of covered and non-covered concepts recalled. As the table indicates, concepts covered with metaphor were generally recalled more often than when they were explained literally, without elaboration. But then why didn't the texts with metaphor do better in general? We can only speculate that those concepts not covered with metaphor in the metaphor versions were not recalled as often, supporting our concern that readers may be selectively attending more to concepts with metaphor and ignoring those not covered with metaphor. This observation needs to be further tested. do we want to compare all the other concepts in both versions to see if the same trends appear?

To eliminate our anticipated problems with coverage and non-coverage, we sometimes found ourselves simply omitting important target concepts when they did not map on. For example, in writing the extended version of the electricity text, we compared electrical current to crowd behavior. In general, this metaphor was useful in that it explained the concepts of voltage and resistance. However, we could not easily use crowd behavior to explain the function of batteries. We therefore eliminated this concept from the electricity passage. In constructing the immunity passage, we began with the war metaphor, creating a framework of enemy soldiers and a defense army, and selecting only target concepts that served as good candidates for this metaphor. Of course, actual textbook writers would be much more constrained in that they often lack control over the content they are required to include in a given section of a textbook. It seemed ironic that our prescriptive use of metaphor sometimes controlled what we could teach instead of helping us teach those concepts which we thought were important.

Although there were instances in which we did find well-structured, apt bases from which to construct metaphors, we sometimes could not use them because of audience considerations. For example, water pressure seemed like an accurate and comprehensive way to explain the concepts of electrical current, yet we had to discard this idea because we knew that many of our students were no more likely to understand water pressure than water current. In deciding to use real life, common scenarios (anthropomorphic rather than technical bases), we realized that we were making assumptions about the kind and structure of base

systems to which readers would be likely to relate. The role and relationship of reader knowledge and learning from certain types of metaphorical bases should be explored in future research.

Once we had located appropriate metaphors, we faced additional choices in how to weave the metaphor into our target information. We began to realize that this staging factor could also affect our readers. Two methods of integrating metaphor into text became apparent. The first was to thoroughly integrate the base into the target material, simply implying the comparison, while the second method involved "tacking on" the comparison in an explicit way. For example, one can simply say that "Pathogen enemies invade the body." This approach requires readers to make a number of inferences in order to fully realize the compared relations between pathogens and the body and soldiers and a foreign country. A second approach makes the metaphor more explicit and visible, and requires less processing: "Pathogens are harmful organisms that enter the human body. You might think of them as enemy soldiers, invading your homeland." This second approach is wordier and also doubles and perhaps reinforces the concept already introduced. It serves to separate the target and base information. We are not entirely certain how these presentational approaches might have affected our readers, but certainly our approach was not always consistent throughout the texts. Although the second, more explicit way of staging the metaphor seemed lengthier, overexplicit and perhaps patronizing, we decided to take this approach when possible for the sake of consistency. However, even some of the younger technical writers in our first study did report that this approach was "condescending," especially in the extended versions. For example one writer remarked, "It forces the reader to read more copy than is required," while other writers remarked that the extended metaphor was "just too simple and patronizing" and "nearly takes over the text."

At least in the ad hoc versions readers had some breathing space in which they were not being spoon fed metaphors after every concept. This made us wonder whether there are stylistic advantages in the ad hoc approach. In the extended versions, the base can overwhelm the target when it keeps emerging; moreover, too much elaboration and explanation can be insulting. Future research on metaphor would do well to investigate the presentational strategies involved in staging metaphors, whether extended or ad hoc, and the effects of this staging on readers.

Our attempt to compose these texts made it clear that the opportunity to use metaphor and the kinds of metaphor one chooses to use are highly constrained by the type and organization of concepts in the topics one hopes to teach. Some science texts are concerned with teaching a string of definitions or concepts that are only loosely related, while others present complex networks of operations within whole systems, e.g., the digestive system, the solar system, etc. The former kind of information seems well suited to very brief, ad hoc metaphors-- sometimes just a word long. For example, an elementary chapter on astronomy might introduce the following terms: comet, planet, solar eclipse. In our review of elementary and secondary science texts, we found an abundance of short metaphors shot in quickly to explain concepts such as these. Consider for example, Saturn's "rings" or the "tail" of a comet. Our topics were quite different however, in that each portrayed a well integrated system of relations and operations rather than a discrete set of definitions. We therefore required lengthier metaphors that could convey relations between objects in a system. For topics such as these, apt metaphors are difficult to find, as we have already suggested.

There were structural differences among our topics as well. The perception topic was extremely procedural, abstract and followed a definite sequence of ordered operations. The structure of this topic lent itself to any number of human scenarios-- narratives involving a series of events in which people act on input, make decisions and produce output. In contrast, the immunity topic was far more complex, involving not an ordered string of procedures-- for which many metaphors might be available, but a hierarchical set features and operations that interacted in critical ways. It was therefore important to find a very large and flexible base in which we could create similar sets of relationships. Fewer metaphors came to mind for this topic.

The number of concepts in a topic is also critical and can also effect the type of base available; the shorter topics could employ more specific and well-structured metaphorical bases. Large topics require larger and more loosely structured bases that can be adapted to the complex number and type of relationships in the topic. We cannot discount the possible effects of topic structure and length on type of metaphor, and the metaphors' subsequent effects on readers' recall of the topic.

Conclusions. In designing metaphors for this particular study, we realized the difficulty of using models and theories of metaphor as specifications for writing. The models themselves say little about the conditions under which metaphors are used and the wide range of approaches involved in choosing and integrating metaphors into text. We were reminded here of the failure to implement readability formulas. Although researchers can examine the properties of successful texts, Duffy (19xx) has shown that when these properties are operationalized into a set of rules for writing, they can subvert the writer's goal to produce readable text and can actually lead to incoherent text. Likewise, our attempt to use comprehensive metaphor often resulted in strained or over-elaborated explanations which may have affected our readers in negative ways. In order to help writers, we need to recognize the importance of these context-specific, rhetorical factors that are often ignored in models of metaphor, for example the size and structure of the topic, the ways in which metaphor can be woven into a topic, the reader's familiarity with and preference for certain kinds of base information from which writers construct metaphors. We can only conclude that metaphor itself will interact with the material at hand and with the particular reader in unpredictable ways that are not necessarily taken into account by the models and theories we have discussed.

Future research will need to investigate these important variables and their effects on learning from metaphor. As writers, we had to respond to all of the factors mentioned above and in doing so we recognized that the metaphors we employed in one text differed in many ways from the metaphors we employed in the other texts. Studying the effects of texts that use metaphor presents a difficult challenge in that we cannot assume that the texts we study are equally comparable.

The Effect of Topic Knowledge In Editing

It is widely believed that if a person is familiar with a topic, that very familiarity may make it difficult for the person to explain the topic to another. Further, research on revision has suggested a mechanism whereby knowledge might make it more difficult to write clearly. Hayes et al. (1987) proposed that "...evaluation is best viewed as an extension of the familiar process of reading for comprehension" (p 202). In particular, they propose that writers

detect problems in text by attending to failures of their own comprehension problems. Any special condition of the writer that makes comprehension easier, such as familiarity with the subject matter, will make it more difficult for the writer to identify parts of the text that might confuse the intended reader. Hayes, Schriver, Spilka and Blaustein (1986) have called this "the knowledge effect" and have examined it in a series of studies.

Study 1

In study 1, 88 undergraduates were asked to read four two-page texts, a clear and an unclear version on each of two topics, autism and statistics. The clear autism and the unclear statistics texts were naturally occurring texts. The statistics text was rewritten to be clear and the autism text to be unclear. For example, in the autism text, the phrase, "Autistic children look normal" was replaced by "Autistic patients in childhood appear asymptomatic". Observations of Study 2, to be described below, confirmed that the clear texts did indeed present fewer comprehension problems to the readers than did the unclear texts. The subjects were then asked to predict reader troubles by underlining those parts which they thought would confuse another student. This will be called the prediction task.

Half of the subjects, the high knowledge group, had read and evaluated a clear version of the text before they read the unclear version. As a result, they had knowledge about the content of the unclear version when they were trying to predict what parts of it would be unclear to other readers. The other half of the subjects, the low knowledge group, had not read the clear version and therefore had little knowledge of the content of the unclear version when they were making their predictions. The results, shown in Figure 8, indicate that knowledge of the text had a strong effect in reducing the numbers of text problems predicted. The low knowledge group identified about 60% more text items as being problematic than did the high knowledge group. This difference between the two groups is highly reliable statistically. ($F= 13.7$; $df=1,88;p<.001$)

Hayes et al (1986) also explored the effect of a short delay between acquiring topic information and evaluating the unclear text. They reasoned that if the subjects information about a topic very recently, coming across that topic in the unclear text might remind

them that they had just learned it and, perhaps, sensitize them to the possibility that their audience does not know it. For the half of the subjects in the delay condition, the subjects read and evaluated a third text in the interval between reading the clear text and reading the unclear text. For the half of the subjects in the no delay condition, the unclear text was presented immediately after the subject had read and evaluated the clear text.

Figure 9 shows that Subjects in the delay condition identified approximately 25% fewer problems in the texts than did subjects in the no delay condition. While this difference is not statistically reliable, it does suggest that delay may intensify the knowledge effect. A subsequent study by Levine (1987) has shown that a delay of one week resulted in an additional reduction by 25% of the number of text problems predicted.

Study 1 showed that the low knowledge group predicted more text problems than the high knowledge group. However, it might be that the predictions of the low knowledge group were poorer in quality than those of the high knowledge group. That is, the low knowledge subjects may have been less successful than the high knowledge subjects in predicting problems that readers actually have. Study 2 was conducted to answer the quality question by determining what undergraduates actually found confusing about the unclear texts.

Study 2

In Study 2, 20 undergraduates were asked to read the unclear texts sentence by sentence and explain the meaning of each one as if to another student who had not read the text. This will be called the "explanation task". The purpose of the explanation task was to learn what Freshmen actually understood in the texts. If subjects overlooked any points or explained them ambiguously, the experimenter questioned them until it was clear whether the point was understood or not. Half of the subjects read a clear version of the text before they attempted to explain the unclear version. These subjects were comparable to the high knowledge subjects in Study 1. The other half of the subjects, who did not read the clear version of the text before attempting to explain the unclear text, and thus were comparable to the low knowledge subjects in Study 1.

The data of Study 2 served two functions in the investigation. First, They identified the aspects of the unclear texts that low knowledge readers actually have difficulty in understanding. These results together with the subjects' predictions in Study 1 were used to perform a signal detection analysis to compare the ability of subjects in Study 1 to predict readers' comprehension problems. Figure 10 shows d-prime values for high and low knowledge subjects.

Second, the data of Study 2 indicated which items of the unclear text were understood better by high knowledge than low knowledge subjects. This was of interest because a knowledge effect should apply only to those items for which prior reading of the clear text improved the subject's understanding. The items in the unclear texts were divided into three categories: the "positive learning" category, consisting of the 31 items which high knowledge subjects understood better than low knowledge subjects; the "zero learning" category, consisting of the 66 items which were equally well understood by high and low knowledge subjects; and the "negative learning" category, consisting of 15 items better understood by
the low knowledge than the high knowledge subjects.

Figure 11 shows the percentage of items in each category for which more low knowledge subjects predicted reader troubles than did high knowledge subjects. These results suggest that the "knowledge effect" really is an effect of knowledge.

References

Aristotle. (1954) Rhetoric. Translated by W. Rhys Roberts. New York: The Modern Library.

Ballay, J., Graham, K., Hayes, J. R. & Fallside, D. C.

Bereiter, C. and Scardamalia, M. (1987) The Psychology of Written Composition. Hillsdale, N. J.:Erlbaum.

Carey, L., Flower, L., Hayes, J R., Schriver, K., and Haas, C. (1987) Differences in Writers' Initial Task Representations. Office of Naval Research Technical Report #2. Pittsburgh, Pa.: Carnegie Mellon University.

Carroll, J. M. & Thomas, J. (1980). Metaphor and the cognitive representation of computing systems. IBM Watson Research Report, 1-13. Yorktown Heights, NY: IBM Research Center.

Duffy, T., & Kabance, P. (1982) Testing a readable writing approach to text revision. Journal of Educational Psychology 74, No. 5: 733-748.

Flower, L. S., and Hayes, J. R. (1984) Images, plans, and prose:The representation of meaning in writing. Written Communication, 1, 120-160.

Gentner, D. (1980). The structure of analogical models in science. ONR Technical Report Number 4451. Bolt Beraneck and Newman, July.

Gentner, D. (1982). Are scientific analogies metaphors? In D. Miall (Ed.), Metaphor: problems and perspective. Brighton, Sussex, England: Harvester Press Ltd.

Gick, M. L., & Holyoak, K. J. (1980). Analogic problem solving. Cognitive psyhology, 12, 306-355.

Hayes, C. (1987) Planning in the Machining Domain: Using Goal Interactions to Guide Search. Unpublished Masters Thesis. Department of Computer Science, Carnegie Mellon University: Pittsburgh, Pa.

Hayes, J. R. and Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg and E. Steinberg (Eds.), Cognitive Processes in Writing: An Interdisciplinary Approach. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Hayes, J. R., Schriver, K. A., Spilka, R., and Blaustein, A. (1986). If its clear to me it must be clear to them. Paper presented at the Conference on College Composition and Communication, New Orleans.

Hayes, J. R., Flower, L. S., Schriver, K. A., Stratman, J., and Carey, L. (1987) Cognitive Processes in Revision In S. Rosenberg (Ed.) Advances in Psycholinguistics, Volume II: Reading, Writing, and Language Processing. Cambridge University Press, Cambridge, England.

Hayes-Roth, B., and Hayes-Roth, F. A Cognitive Model of Planning. Cognitive Science, 1979, 3, 275-310.

Kaufer, D., Hayes, J. R., and Flower, L. S. (1986). Composing Written Sentences. Research in the Teaching of English. 20, 121-140.

Kittay, E. F. (1987). Metaphor: Its cognitive force and linguistic structure . New York: Oxford University Press.

Kohler, W. (1940) Dynamics in Psychology. New York: Liveright.

Lakoff, G. & Johnson, M. (1980). Metaphors we live by. Chicago: Chicago University Press.

Leatherdale, W. H. (1974). The role of analogy, model and metaphor in science. Amsterdam: North-Holland.

Levine, L. (1987). Personal communication.

Mac Cormac, E. (1985). A cognitive Theory of Metaphor. Cambridge: MIT Press.

McDermott, D. (1978) Planning and Action. Cognitive Science, 2, 71-109.

Miller, G. A., Galanter, E., and Pribram, K. (1960) Plans and the Structure of Behavior. New York:Holt, Rinehart, and Winston.

Newell. A., and Simon, H. A. (1972) Human Problem Solving. Englewood Cliffs, N. J.:Prentice-Hall.

Reitman, W. R. (1965) Cognition and Thought. New York: Wiley.

Reddy, R. and Newell, A. (1974) Knowledge and It's Representation in a Speech Understanding System. In L. W. Gregg (Ed.) Knowledge and Cognition, Lawrence Erlbaum Assoc., Washington, D. C.

Sacerdoti, E. D. (1974) Planning in a Hierarchy of Planning Spaces. Artificial Intelligence, 5, 115-135.

Scardamalia, M., Bereiter, C., & Steinbach, R. (1984) Teachability of reflective proceses in written composition. *Cognitive Science* 8 (2): 173-190.

Schriver, K. A. (1986). Designing computer documentation: A review of the literature. Communication Design Center *Technical Report No. 31*, Pittsburgh, PA: Carnegie Mellon University.

Schriver, K. A. (1987). Teaching writers to anticipate the reader's needs: empirically based instruction. Doctoral dissertation. Pittsburgh, PA: Carnegie Mellon University.

Selfridge, O. G. (1959) Pandemonium, a paradigm for learning. In *The mechanization of thought processes.* London: H. M. Stationary Office.

Shakespeare.,W. (1973) The complete works of Shakespeare. C. Hardin & D. Bevington (Eds.) Glenview, ILL: Scott, Foresman and co.

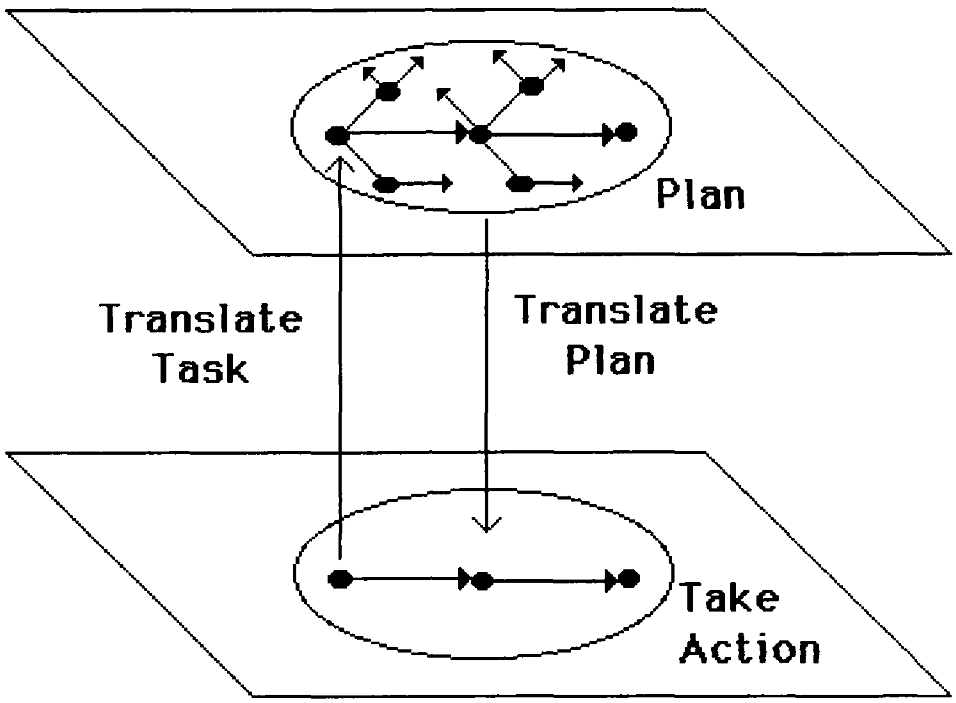
Spiro, R. J., Feltovich, P. J., Coulson, R. L, & Anderson, D. K. (1987). Multiple Analogies for Complex Concepts: Antidotes for Analogy-Induced Misconception in Advanced Knowledge Acquisition. In S. Vosniadou & A, Ortony (Eds.) Similarity and Analogical Reasoning Cambridge: Cambridge University Press.

Stefic, M. (1981) Planning With Constraints. Artificial Intelligence, 16, 111-140.

Stefic, M. (1981) Planning and Meta-Planning (MOLGEN: Part 2) Artificial Intelligence, 16, 141-170.

Swaney, J. H., Janik, C. J., Bond, S. J., and Hayes, J. R. (1981). Editing for comprehension: Improving the process through reading protocols. Carnegie Mellon University, Document Design Project Technical Report #14.

Planning Environment



Action Environment

Figure 1. The relation between planning and action.

Topic A Immunity	Topic B Electricity	Topic C Perception
No Metaphor ("P")	No Metaphor ("P")	No Metaphor ("P")
Ad Hoc ("AH")	Ad Hoc ("AH")	Ad Hoc ("AH")
Extended ("E")	Extended ("E")	Extended ("E")

Figure 2 Design of the Nine Instructional Texts: Topics and Versions

	Plain	Mixed	Extended
Pathogens	0.82	1.12	1.06
Electricity	0.76	0.88	1.35
Perception	0.88	1.41	0.71
Mean	0.82	1.14	1.04

Figure 3. Mean rankings of appropriateness for audience.

	Plain	Mixed	Extended
Pathogens (N = 34)	12.16	13.20	10.14
Electricity (N = 7)	5.80	4.28	3.83
Perception (N = 13)	4.28	3.50	4.20
Mean	7.41	6.99	6.06

Figure 4. Mean numbers of concepts recalled.

	Plain	Mixed	Extended
Pathogens (N = 27)	6.16	7.40	6.42
Electricity (N = 5)	4.00	2.71	2.50
Perception (N = 12)	3.42	2.50	3.60
Mean	4.53	4.20	4.17

Figure 5. Mean number of "hard" concepts recalled.

	Plain	Mixed	Extended
Pathogens	4	12	14
Electricity	2	4	4
Perception	5	4	2
Total	11	20	20

Figure 6. Total number of errors in each condition.

	non-covered (version P)	covered (version AH)
	plain	ad hoc
immunity	3.5	5.6
electricity	1.7	3.2
perception	2.3	1.9
Sum	7.5	10.7

Figure 7. Mean number of covered and non-covered concepts recalled from the Ad Hoc and Plain Versions.

	Knowledge	
	Low	High
Autism	267	161
Graph	127	83
Total	394	244

Figure 8. Numbers of items underlined.

	No Delay	Delay
Autism	87	74
Graph	53	30
Total	140	104

Figure 9. The effect of delay on predictions.

	Knowledge	
	High	low
Autism	.65	.80
Graph	.32	.56

Figure 10. Effect of knowledge on D-prime scores.

	Positive Learning N = 31	Zero Learning N = 66	Negative Learning N = 15
Percent of items for which more low Knowledge Ss predicted difficulties	61.3%	48.5%	46.7%

Figure 11. Effect of knowledge on predictions of difficulty.